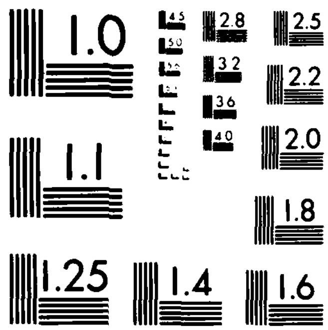
AD-A149 568 THE INEFFICIENCY OF LEAST SQUARES: EXTENSIONS OF KANTOROVICH INEQUALITY(U) PITTSBURGH UNIV PA CENTER FOR MULTIVARIATE ANALYSIS C R RAO NOV 84 TR-84-47 1/1

UNCLASSIFIED AFOSR-TR-84-1199 F49620-82-K-0001 F/G 12/1 NL

END

FILMED

DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963 A

AFOSR-TR- 84-1199

THE INEFFICIENCY OF LEAST SQUARES: EXTENSIONS OF KANTOROVICH INEQUALITY

C. Radhakrishna Rao
University of Pittsburgh

F49620-82-K-0001

Center for Multivariate Analysis

University of Pittsburgh

THE INEFFICIENCY OF LEAST SQUARES:
EXTENSIONS OF KANTOROVICH INEQUALITY

C. Radhakrishna Rao
University of Pittsburgh

F49620-82-K-0001

November 1984

Technical Report No. 84-47

Center for Multivariate Analysis
Fifth Floor, Thackeray Hall
University of Pittsburgh
Pittsburgh, PA 15260

This work was supported by the Air Force Office of Scientific Research. Reproduction in whole or in part is permitted for any purpose of the United States Government.

THE INEFFICIENCY OF LEAST SQUARES: EXTENSIONS OF KANTOROVICH INEQUALITY

C. Radhakrishna Rao

ABSTRACT

Four different measures of inefficiency of the simple least squares estimator in the general Gauss-Markoff model are considered. Previous work on the bounds to some of these measures is briefly reviewed and new bounds are obtained for a particular measure.

Keywords: Inefficiency of least squares, Kantorovich inequality.

1. INTRODUCTION

Let us consider the usual Gauss-Markoff model

$$Y = X\beta + \varepsilon,\ E(\varepsilon) = 0,\ D(\varepsilon) = \sigma^2 V \tag{1.1}$$

where Y and ε are n-vectors, β is an m-vector, X is an $n \times m$ matrix and $\sigma^2 V$, the dispersion matrix of ε, is positive definite. In practice V may be unknown in which case the estimate of β is computed by choosing an apriori dispersion matrix V_a in the place of V. A number of authors have investigated the loss of information in the estimation of β resulting from a wrong choice of V. See for instance, Bloomfield and Watson (1975), Khatri and Rao (1981,1982), Knott (1975) and Styan (1983). The object of the present paper is to review some of the earlier results and provide a generalization of a recent result by Styan (1983).

There is no loss of generality in assuming $V_a = I$, for the problem associated with V_a and V, when V_a is positive definite, is the same as that with I and $V_a^{-1/2}\ V\ V_a^{-1/2}$ for purposes of the present investigation. Also, we consider the basic parameter as $X\beta$ and study the inefficiency of its estimation due to a wrong choice of V. In such a case, we can, without loss of generality, consider X to be of full rank with its column vectors as orthonormal. Then the simple least squares estimator and the BLUE of $X\beta$ are

$$XX'Y \text{ and } X(X'V^{-1}X)^{-1}X'V^{-1}Y \tag{1.2}$$

with the dispersion matrices (apart from the multiplying constant σ^2)

$$XX'VXX' \text{ and } X(X'V^{-1}X)^{-1}X' . \tag{1.3}$$

If $V = P \Lambda P'$ is the spectral decomposition of V, where P is an orthogonal matrix and Λ is a diagonal matrix of eigenvalues of V, then the matrices in (1.3) can be written as

$$PUU' \Lambda UU'P' \quad \text{and} \quad PU(U'\Lambda^{-1}U)^{-1}U'P' \tag{1.4}$$

where $U = P'X$ and hence $U'U = I$. In the next section we define a number of measures of inefficiency based on a comparison of the matrices in (1.3) for (1.4) and determine their lower and upper bounds.

2. MEASURES OF INEFFICIENCY AND THEIR BOUNDS

By construction, the difference between the first and second matrices in (1.4) is non-negative definite and the magnitude of the difference can be judged by the magnitudes of the proper nonzero eigenvalues (Rao and Mitra, 1971, 124-126) of the first matrix with respect to the second. If θ is a proper nonzero eigenvalue, then

$$PUU'\Lambda UU'P'x = \theta PU(U'\Lambda^{-1}U)^{-1}U'P'x \tag{2.1}$$

where the left and right hand sides of (2.1) do not individually vanish. In such a case, multiplying both sides of (2.1) by U'P' and writing $y = U'P'x$, we have

$$U'\Lambda Uy = \theta(U'\Lambda^{-1}U)^{-1}y \tag{2.2}$$

so that θ is a root of the determinental equation

$$|U'\Lambda U - \theta(U'\Lambda^{-1}U)^{-1}| = 0. \tag{2.3}$$

We may choose any increasing function of the roots $\theta_1,\ldots,\theta_m$ of (2.3) as a measure of inefficiency, such as $\theta_1 \ldots \theta_m$ or $\theta_1 + \ldots + \theta_m$. Bloomfield and Watson (1975) and Knott (1975) have established the bounds

$$1 \le \prod_1^m \theta_i \le \sum_{i=1}^{s} \frac{(\lambda_i + \lambda_{n-i+1})^2}{4\lambda_i\lambda_{n-i+1}} \tag{2.4}$$

where $\lambda_1 \ge \lambda_2 \ldots \ge \lambda_n$ are the diagonal elements of Λ (i.e., the eigenvalues of V) and $s = \min(m, n-m)$. Khatri and Rao (1981) established that

$$m \le \sum_1^m \theta_i \le \sum_1^s \frac{(\lambda_i + \lambda_{n-i+1})^2}{4\lambda_i\lambda_{n-i+1}} + t \tag{2.5}$$

where $s = \min(m, n-m)$, and $t = 0$ if $s = m$ and $t = 2m-n$ if $s = n-m$.

Puntanen (1982) suggested the use of

$$\mathrm{tr}(PUU'\Lambda UU'P' - PU(U'\Lambda^{-1}U)^{-1}U'P') \tag{2.6}$$

as a measure of inefficiency. The expression (2.6) reduces to

$$\mathrm{tr}(U'\Lambda U - (U'\Lambda^{-1}U)^{-1}). \tag{2.7}$$

In the special case when m = 1, Styan (1983) showed that

$$0 \le U'\Lambda U - (U'\Lambda^{-1}U)^{-1} \le (\sqrt{\lambda_1} - \sqrt{\lambda_n})^2. \tag{2.8}$$

We provide the following generalization for higher values of m.

Theorem. For general m

$$0 \le \mathrm{tr}(U'\Lambda U - (U'\Lambda^{-1}U)^{-1}) \le \sum_{i=1}^{s} (\sqrt{\lambda_i} - \sqrt{\lambda_{n-i+1}})^2 \tag{2.9}$$

where $s = \min(m, n-m)$.

Proof. We find the stationary values of (2.7) subject to the condition $U'U = I$. Introducing a symmetric matrix A of Lagrangian multipliers we consider the expression

$$\mathrm{tr}\, U'\Lambda U - \mathrm{tr}(U'\Lambda^{-1}U)^{-1} - \mathrm{tr}\, A(U'U - I). \tag{2.10}$$

Taking derivatives of (2.10) with respect to the elements of U (see Rao, 1984) and equating to zero, we have

$$U'\Lambda + (U'\Lambda^{-1}U)^{-2}\, U'\Lambda^{-1} = AU' \tag{2.11}$$

which gives

$$U'\Lambda U + (U'\Lambda^{-1}U)^{-1} = A. \tag{2.12}$$

Then the equation (2.11) reduces to

$$U'\Lambda + (U'\Lambda^{-1}U)^{-2}\, U'\Lambda^{-1} = (U'\Lambda U + (U'\Lambda^{-1}U)^{-1})U'. \tag{2.13}$$

Multiplying both sides of (2.13) from the right by ΛU

$$U'\Lambda^2 U + (U'\Lambda^{-1}U)^{-1} = (U'\Lambda U)^2 + (U'\Lambda^{-1}U)^{-1}(U'\Lambda U) \tag{2.14}$$

which shows that the last term in (2.14) is symmetric or the matrices $U'\Lambda U$ and $U'\Lambda^{-1}U$ commute. Then, there exists an orthogonal matrix Q such that

$$U'\Lambda U = Q\,E Q' \text{ and } U'\Lambda^{-1}U = Q\,\Delta\,Q' \tag{2.15}$$

where Δ and E are diagonal matrices with diagonal elements, say, $d_1,\ldots,d_m$ and $e_1,\ldots,e_m$. Writing $W = UQ$, the equation (2.13) becomes

$$\Lambda W + \Lambda^{-1} W \Delta^{-2} = W(E + \Delta^{-1}). \tag{2.16}$$

Let $(w_1,\ldots,w_n)'$ be the j-th column of W. Then

$$\lambda_i w_i + \lambda_i^{-1}\ w_i\, d_j^{-2} = w_i(e_j + d_j^{-1}),\ i = 1,\ldots,n \tag{2.17}$$

which shows that at most two values of w_i can be nonzero. If w_r and w_s are non-zero, then

$$\lambda + \lambda^{-1} d_j^{-2} = (e_j + d_j^{-1}) \tag{2.18}$$

has two roots λ_r and λ_s, and it is seen that

$$e_j - d_j^{-1} = (\sqrt{\lambda_r} - \sqrt{\lambda_s}\,)^2. \tag{2.19}$$

If only one w_i is non-zero, then

$$e_j - d_j^{-1} = 0. \tag{2.20}$$

The expression we have to maximize is

$$\begin{aligned} \operatorname{tr}(U'\Lambda U - (U'\Lambda^{-1}U)^{-1}) &= \operatorname{tr}(E - \Delta^{-1}) \\ &= \sum_1^m (e_j - d_j^{-1}) \end{aligned} \tag{2.21}$$

where each term in (2.21) has the value zero as in (2.20) or a value of the type $(\sqrt{\lambda_r} - \sqrt{\lambda_s}\,)^2$ as in (2.19). Using arguments similar to those in Bloomfield and Watson (1975) and Knott (1975), we find the maximum of (2.7) is

$$\sum_{i=1}^{s} (\sqrt{\lambda_i} - \sqrt{\lambda_{n-i+1}}\,)^2 \tag{2.22}$$

where $s = \min(m, n-m)$, which proves the required result.

Remark 1. In terms of the original matrices X and V where X need not be assumed to have orthonormal columns, the matrices in (1.3) can be written as

$$P_X V P_X \text{ and } P_X(P_X V^{-1} P_X)^- P_X \tag{2.23}$$

where $P_X = X(X'X)^- X'$, the projection operator on the space generated by the columns of X and $(\cdot)^-$ denotes any generalized inverse. Then the result (2.9) of the Theorem can be written as

$$0 \leq \operatorname{tr}(P_X V P_X - P_X(P_X V^{-1} P_X)^- P_X) \leq \sum_{i=1}^{s} (\sqrt{\lambda_i} - \sqrt{\lambda_{n-i+1}})^2. \tag{2.24}$$

where λ_i are the eigenvalues of V. Let x_i be the corresponding eigenvectors and denote

$$\xi_i = \left[\frac{\sqrt{\lambda_i}}{\sqrt{\lambda_i} + \sqrt{\lambda_{n-i+1}}}\right]^{\frac{1}{2}} x_i \pm \left[\frac{\sqrt{\lambda_{n-i+1}}}{\sqrt{\lambda_i} + \sqrt{\lambda_{n-i+1}}}\right]^{\frac{1}{2}} x_{n-i+1} \tag{2.25}$$

It is seen that the upper bound in (2.24) is attained when the columns of X are generated by the vectors $\xi_1,\ldots,\xi_s$ and some x_i vectors orthogonal to $\xi_1,\ldots,\xi_s$.

Remark 2. When $m = 1$, simple proofs are available for the inequalities (2.4) and (2.9) as given by Styan (1983).

3. ANOTHER MEASURE OF INEFFICIENCY

In a paper presented at the Fifth Berkeley Symposium in 1965, the author showed that there is no loss of information in estimation by simple least squares if V and X satisfy the condition $X'VZ = 0$ where Z is any matrix with maximum rank such that $X'Z = 0$ (see Rao, 1967). The equivalent condition $P_X V = VP_X$ was given by Zyskind (1967). The condition $X'VZ = 0 = P_X VZ$ is equivalent to

$$\begin{aligned} 0 &= P_X VZ(Z'Z)^- Z'VP_X = P_X V(I - P_X) V P_X \\ &= P_X V^2 P_X - (P_X V P_X)(P_X V P_X). \end{aligned} \tag{3.1}$$

Bloomfield and Watson (1975) considered (3.1) as a measure of inefficiency and showed that

$$0 \le \mathrm{tr}(P_X V^2 P_X - P_X V P_X V P_X) \le \sum_{i=1}^{s} (\lambda_i - \lambda_{n+i-1})^2 .$$

4. REFERENCES

[1] BLOOMFIELD, P. and WATSON, G.S. (1975). The inefficiency of least squares. Biometrika 62, 121-128.

[2] KHATRI, C.G. and RAO, C.R. (1981). Some extensions of the Kantorovich inequality and statistical applications. J. Multivariate Analysis 11, 498-505.

[3] KHATRI, C.G. and RAO, C.R. (1982). Some generalizations of Kantorovich inequality. Sankhya A 44, 91-102.

[4] KNOTT, M. (1975). On the minimum efficiency of least squares. Biometrika 62, 129-132.

[5] PUNTANEN SIMO (1982). Personal communication.

[6] RAO, C.R. (1967). Least squares theory using an estimated dispersion matrix and its application to measurement of signals. Proc. Fifth Berkeley Symp. 355-373.

[7] RAO, C.R. and MITRA, S.K. (1971). Ge' ,ized Inverse of Matrices and its Applications, Wiley, New York

[8] RAO, C. RADHAKRISHNA (1984). Matrix der. ıves: Applications in Statistics. Encyclopedia of Statistical Sciences, (Eds. S. Kotz and N. L. Johnson), Wiley.

[9] STYAN, G.P.H. (1983). On some inequalities associated with ordinary least squares and the Kantorovich inequality. In Festschrift for Eino Haikala on his Seventieth Birthday, University of Tampere, 158-166.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) 84-47	5. MONITORING ORGANIZATION REPORT NUMBER(S) AFOSR-TR- 84-1199
6a. NAME OF PERFORMING ORGANIZATION University of Pittsburgh — 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION Air Force Office of Scientific Research
6c. ADDRESS (City, State and ZIP Code) Center for Multivariate Analysis Fifth Floor, Thackeray Hall Pittsburgh PA 15260	7b. ADDRESS (City, State and ZIP Code) Directorate of Mathematical & Information Sciences, Bolling AFB DC 20332-6448
8a. NAME OF FUNDING/SPONSORING ORGANIZATION AFOSR — 8b. OFFICE SYMBOL (If applicable) NM	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER F49620-82-K-0001
8c. ADDRESS (City, State and ZIP Code) Bolling AFB DC 20332-6448	10. SOURCE OF FUNDING NOS. PROGRAM ELEMENT NO. 61102F; PROJECT NO. 2304; TASK NO. A5; WORK UNIT NO.

11. TITLE (Include Security Classification)
THE INEFFICIENCY OF LEAST SQUARES: EXTENSIONS OF KANTOROVICH INEQUALITY

12. PERSONAL AUTHOR(S)
C. Radhakrishna Rao

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Yr., Mo., Day)	15. PAGE COUNT
Technical	FROM ______ TO ______	NOV 84	8

16. SUPPLEMENTARY NOTATION

17. COSATI CODES		
FIELD	GROUP	SUB. GR.

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
Inefficiency of least squares; Kantorovich inequality.

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Four different measures of inefficiency of the simple least squares estimator in the general Gauss-Markov model are considered. Previous work on the bounds to some of these measures is briefly reviewed and new bounds are obtained for a particular measure.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS ☐	21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL MAJ Brian W. Woodruff	22b. TELEPHONE NUMBER (Include Area Code) (202) 767- 5027 — 22c. OFFICE SYMBOL NM

85 01 16 104

DD FORM 1473, 83 APR EDITION OF 1 JAN 73 IS OBSOLETE.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

END

FILMED

2-85

DTIC